

# A delay optimal policy for symmetric broadcast channels

Stéphane Musy  
School of Computer and Communication Sciences, EPFL  
CH-1015 Lausanne, Switzerland  
Email: stephane.musy@epfl.ch

**Abstract**—We consider the rate allocation problem when two users (each one associated with one receiver) send packets through a symmetric broadcast channel. Under the assumption that the packet lengths are exponentially distributed, we establish the delay optimality of the Best User Highest Possible Rate (BUHPR) scheduling policy. Such a policy always allocate the whole possible rate to the non-empty queue serving the user with the minimum average packet length. Our analysis can be generalized to the symmetric multiuser broadcast channel.

## I. INTRODUCTION

In multiuser communication systems the problem of allocating resources (such as rate or power) between bursty users, has shown to be essential in the characterization of the optimal values for quality-of-service measures like packet throughput and delay. In [1], [2] and the references therein, the need for a cross-layer approach between information theory and network theory to resource allocation has been pointed out. Consequently, the rate allocation is formulated as a scheduling problem where the rate of each user is allocated from the channel capacity region according to the state of their queue.

In [3], [4], the network capacity region is defined as the set of all packet arrival rates for which the queue lengths can be stabilized. A scheduling policy that achieves the network capacity region is called *throughput optimal*. For Gaussian multiaccess channels and Gaussian broadcast channels, policies that give priority to the longest queues<sup>1</sup> such as Longest Queue Highest Possible Rate (LQHPR) are shown to be throughput optimal [2], [4]. However, stability in a queueing system does not indicate how large the queue sizes can be. To minimize the average packet delay, it is necessary to keep the queue lengths as short as possible. A policy that minimizes the average queue sizes is called *delay optimal*. The authors of [2], establish that for the Gaussian multiple access channel with symmetric setting and when the packet arrival times follow a Poisson distribution with equal rates, the LQHPR policy is also delay optimal.

In this paper, we focus on the symmetric broadcast channel and show that delay optimal policies should allocate rates at the corner points of the capacity region. This combined with

a well known result of scheduling theory, allows us to prove that the BUHPR policy is delay optimal under the restriction of exponential packet lengths. A particular case, where the packet arrivals follow a Poisson distribution, is treated in [5].

The next section describes the queueing systems considered in this paper and formulate the scheduling problem. In Section III, symmetric broadcast channels are defined and we show that any delay optimal policies should allocate rates at the corner points of the capacity region. Then, in Section IV, we show that a BUHPR scheduling policy is delay optimal for the system treated in this paper. Finally, Section V, provides some concluding remarks.

## II. SYSTEM MODEL

The system is composed of two users transmitting independent messages through a symmetric broadcast channel (SBC). Each user generates a data stream of packets according to an arrival process  $\{A_i(t), i = 1, 2\}$ , with rates  $\{\lambda_i(t), i = 1, 2\}$ . No statistical restriction is imposed on the arrival processes, particularly  $\lambda_1$  and  $\lambda_2$  can vary with the time. The packet lengths are i.i.d. random variables exponentially distributed with mean  $\{\mu_i^{-1} < \infty, i = 1, 2\}$ , further independent of the arrival processes.

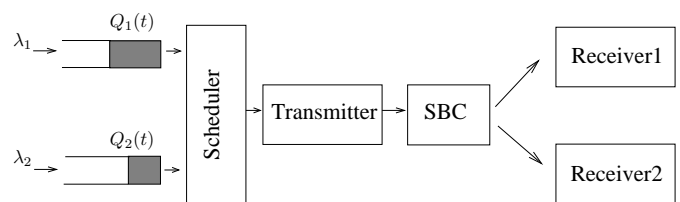


Fig. 1. System Model

Packets from user  $i$  enter queue  $i$  and wait until they are served to receiver  $i$ . The user's queues are assumed to have infinite capacity. Let us denote by  $Q_i(t)$  the number of packets in queue  $i$  at time  $t$ . At each decision epoch (arrival and service completion epoch), the scheduler allocates a rate vector within the capacity region  $\mathcal{R}_{SBC}$  defined in the next section.<sup>2</sup>

The admissible scheduling policies can use information from the number of packets and the total unfinished work (untransmitted bits) in the queues, the previous decisions as

The work presented in this paper was partially supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

<sup>1</sup>They are examples of Maximum Weight Matching Scheduling (MWMS)

<sup>2</sup>We assume that reliable communication is feasible for any rates within the capacity region.

well as the future packet arrival times.<sup>3</sup> But, they do not know the future packet lengths (service requirements). The block diagram of the system model is illustrated in Fig. 1. Here, we are interested in delay optimal policies in the sense that they minimize the cost function  $\sum_{i=1}^2 E[Q_i(t)]$ , over all admissible scheduling policies.

### III. SYMMETRIC BROADCAST CHANNEL

Let  $x$  denote the transmitter output signal, and  $y$  (resp.  $z$ ) denote the receive signal at receiver 1 (resp. 2). Two user broadcast channels are memoryless channels defined by the transition probability  $p(y, z|x)$ . Symmetric broadcast channels form a subclass in which the channel statistics of both links (transmitter to receiver 1 and transmitter to receiver 2) are the same.<sup>4</sup> They are special cases of degraded broadcast channels for which the capacity region is well known [6]. For symmetric broadcast channels, the capacity region is given by the triangular shape region (see Fig. 2) defined as

$$\mathcal{R}_{SBC} = \{R_1 + R_2 \leq C, R_1 \geq 0, R_2 \geq 0\},$$

where  $C$  is the maximum achievable rate over one link.

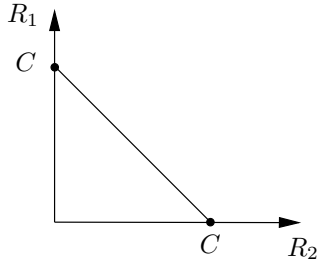


Fig. 2. Example of a symmetric broadcast channel capacity region

Observe that a similar region is obtained for broadcast channels, if we only allow time-sharing coding strategies. One can argue that, since all the rates in  $\mathcal{R}_{SBC}$  can be achieved using time-sharing, the scheduler should allocate the whole rate to one user at a time. But, it is not sure that a policy using rates on the dominant face of the capacity region will not perform better in terms of packet delay. Here, we give a proposition showing that such policies cannot beat a strategy operating at the corner points of the capacity region.

**Proposition 3.1:** For symmetric broadcast channels, a delay optimal scheduling policy should always operate at the corner points of the capacity region  $\mathcal{R}_{SBC}$ .

*Proof:* Let  $V_1$  and  $V_2$  denote the number of untransmitted bits present in the queue of user 1 and 2. First, assume that no more packets arrive in either queue. Without loss of generality (w.l.o.g.) assume that user 1 is served with a rate such that he will clear his queue before user 2. Denote by  $(R_1^*, R_2^*)$ , the rate pair at which both queues are processed during the time that the first queue is emptied, w.l.o.g. we can assume that

this rate pair lies in the dominant face of  $\mathcal{R}_{SBC}$ . Therefore, from a receiver's perspective, the transmission rates seen by the users are

$$\begin{aligned} R_1 &= R_1^* \\ R_2 &= rR_2^* + (1-r)C, \end{aligned}$$

where  $r = \frac{R_2 V_1}{R_1^* V_2} \triangleq \frac{R_2}{R_1^*} q$ , is the fraction of the second user's transmission time, during which the two queues are serviced simultaneously. Note that we have  $r \leq 1$  which implies  $R_1^* \geq qR_2^*$ . Then, we can develop  $R_2$  as follows

$$R_2 = \frac{R_2}{R_1^*} q R_2^* + (1 - \frac{R_2}{R_1^*} q) C,$$

implying

$$R_2 = \frac{R_1^* C}{R_1^* + q(C - R_2^*)}.$$

Since  $(R_1^*, R_2^*)$  is in the dominant face of  $\mathcal{R}_{SBC}$ , we have  $R_2^* = C - R_1^*$ . Thus, we get

$$\begin{aligned} R_2 &= \frac{R_1^* C}{R_1^* (1 + q)} \\ &= \frac{C}{1 + q}, \end{aligned}$$

which is a constant independent of the value of  $R_1^*$ . Hence, the choice  $R_1^* = C$  simultaneously maximizes  $R_1$  and  $R_2$ , and thus minimizes the delay of each packet. Now, removing the assumption that no packets arrive, we see that the waiting time of any new packets can only be increased by not sending previous packets at rates lying at the corner points of  $\mathcal{R}_{SBC}$ . ■

This proposition can be extended to broadcast channels restricted to operate in the time-sharing region. Notice that the LQHP policy used in [2], functions also at the corner points of the Gaussian multiaccess channel capacity region. The delay optimality of strategies operating at the extreme points is emphasized in [5].

### IV. DELAY OPTIMAL POLICY

The result of the previous section (Proposition 3.1), shows that a delay optimal policy should serve one user at a time with the maximum possible rate. Focusing on those policies allows us to formulate the problem as a G/M/1 queue with two classes of customers, where the scheduling policy assign at each decision epoch the entire service rate to one class of user.

This scheduling problem has already been studied, and we know that under the same assumptions as the ones made in Section II, an optimal policy follows a  $\mu c$ -rule. More precisely, in a G/M/1 system with  $K$  classes of customers, define  $\mu_k^{-1}$  as the average service requirement of customers in class  $k$ , we want to address the minimization of the cost function  $\sum_{k=1}^K c_k E[Q_k(t)]$ , where  $c_k$ 's are arbitrary non-negative numbers. It is shown (see, e.g., [7]), that a scheduling policy that always allocates the server to the non-empty queue with highest  $\mu_k c_k$  (the so called  $\mu c$ -rule) minimizes the cost function over all admissible policies, for all  $t \geq 0$ .

<sup>3</sup>This can be an implementation issue. However, we will see that a non-anticipating policy is optimal.

<sup>4</sup>Symmetric Gaussian broadcast channels with a total power constraint can be inserted in this subclass.

Observe that in our setup the packet lengths are the service requirements, and the queue is serviced at a rate equal to  $C$ . Thus using the result about the  $\mu c$ -rule, we are able to establish the delay optimality of the Best User Highest Possible Rate (BUHPR) policy. Indeed, such a policy always allocates the possible rate to the non-empty queue with the highest  $\mu_i$ ,  $i = 1, 2$ . We have the following theorem on the delay optimality of BUHPR policies.

*Theorem 4.1:* Assume that  $\mu_1 \geq \mu_2 \geq 0$ . Then, for symmetric broadcast channels, the BUHPR scheduling policy minimizes the cost function  $\sum_{i=1}^2 E[Q_i(t)]$  over all admissible policies, for all  $t \geq 0$ .

*Proof:* Proposition 3.1 and Theorem 4.3 in [7] with  $\{c_k = 1, k = 1, 2\}$ . ■

Observe that, although admissible policies are able to look at the future arrival times, the BUHPR policy does not need to know them. In the case of broadcast channels constrained in the time-sharing region, a similar result holds by appropriately scaling  $\mu_1$  and  $\mu_2$ . Let  $C_1$  (resp.  $C_2$ ) be the maximum rate achievable by a user sending information to the receiver 1 (resp. 2), then the BUHPR policy will allocate the possible rate to the non-empty queue with the highest  $\mu_i C_i$ ,  $i = 1, 2$ .

## V. REMARKS AND CONCLUSION

For symmetric broadcast channels, we showed the delay optimality of the BUHPR scheduling policy, under the assumption of exponential packet lengths. Since no restrictions are made on the arrival rates, it should be clear that the BUHPR policy is also throughput optimal, at least for the system model of this paper. For the sake of simplicity we have only considered two users, but our analysis can be extended to the symmetric multiuser broadcast channel.

Note that, the BUHPR policy is a preemptive service discipline and we have assumed that any rates within the capacity region can be allocated. Thus it is not sure that a particular coding scheme, achieving these requirements, exist. Nevertheless, the performance of the optimal policies derived in this context, establishes a fundamental bound to the performance of any reliable coding scheme.

## ACKNOWLEDGMENT

The author wishes to thank Emre Telatar for insightful discussions and helpful comments.

## REFERENCES

- [1] E. Telatar and R. G. Gallager, "Combining Queueing Theory with Information Theory for Multiaccess," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 6, pp. 963-969, August 1995.
- [2] E. Yeh and A. Cohen, "Information Theory, Queueing, and Resource Allocation in Multi-user Fading Communications," *Proceedings of the 2004 Conference on Information Sciences and Systems*, Princeton, NJ, March 2004, Invited Paper.
- [3] L. Tassiulas and A. Ephremides, "Stability Properties of Constrained Queueing Systems and Scheduling Policies for Maximum Throughput in Multihop Radio Networks," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 1936-1948, December 1992.
- [4] M. J. Neely, E. Modiano, and C. R. Rohrs, "Power Allocation and Routing in Multibeam Satellites with Time-Varying Channels," *IEEE/ACM Trans. Networking*, vol. 11, no. 1, pp. 138-152, 2003.
- [5] S. Raj and E. Telatar "Queueing and Scheduling in Multiuser Communications," *Winter School on Coding and Information Theory*, Bratislava, February 2005.
- [6] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York, 1990.
- [7] Z. Liu, P. Nain and D. Towsley, "Sample path methods in the control of queues," *Queueing Systems*, vol. 21, pp. 293-335, 1995.